

OIFE

RAW SEQUENCE LISTING

DATE: 07/01/2000

PATENT APPLICATION: US/09/597,920

TIME: 12:40:32

Input Set : A:\05033.app

Output Set: N:\CRF3\07012000\I597920.raw

ENTERED

3 <110> APPLICANT: Samelson, Lawrence E.
 4 Zhang, Weiguo
 6 <120> TITLE OF INVENTION: Compositions and Methods for Identifying and Testing
 7 Tyrosine Kinase Substrates and Their Agonists and
 8 Antagonists
 10 <130> FILE REFERENCE: NIH-05033
 C--> 12 <140> CURRENT APPLICATION NUMBER: US/09/597,920
 C--> 13 <141> CURRENT FILING DATE: 2000-06-19
 15 <150> PRIOR APPLICATION NUMBER: 60/068,690
 16 <151> PRIOR FILING DATE: 1997-12-23
 18 <160> NUMBER OF SEQ ID NOS: 15
 20 <170> SOFTWARE: PatentIn Ver. 2.0
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 1059
 24 <212> TYPE: DNA
 25 <213> ORGANISM: Homo sapiens
 27 <400> SEQUENCE: 1
 28 gactctgccc ttgaggggccc taggggtgca gccagcctgc tccgagctcc cctgcagatg 60
 29 gaggaggcca tcctgggtccc ctgcgtgctg gggctcctgc tgcggccatc ctggccatgt 120
 30 tgatggcact gtgtgtgcac tgccacagac tgccaggctc ctacgacagc acatcctcag 180
 31 atagtttgta tccaaggggc atccagttca aacggcctca caccggttgc ccctggccac 240
 32 ctgcctaccc acctgtcacc tcctaccac ccctgagcca gccagacctg ctccccatcc 300
 33 caagatcccc gcagcccctt gggggctccc accggagccc atcttcccgg cgggattctg 360
 34 atgggtgcaa cagtgtggcg agctacgaga acgaggaacc agcctgtgag gatgcagatg 420
 35 aggatgagga cgactatcac aaccaggct acctggtggt gcttccctgac agcaccctcg 480
 36 ccactagcac tgctgcccac tcagctcctg cactcagcac ccctggcacc cgagacagtg 540
 37 ccttctccat ggagtccatt gatgattacg tgaacgttcc ggagagcggg gagagcgag 600
 38 aagcgtctct ggatggcagc cgggagtatg tgaatgtgct ccaggaactg catcctggag 660
 39 cggtctaacg tgagcctgccc gccctgagtt cccaggagc agaggaagtg gaggaagagg 720
 40 gggctccaga ttacgagaat ctgcaggagc tgaactgagg gcctgtggag gccgagctg 780
 41 tcctggaacc aggccttgcct gggacggctg agctgggagc ctggaagtgg ctctggggtc 840
 42 ctcacatggc gtctgtccct tgctccagcc tgacaacagc ctgagaaatc cccccgtaac 900
 43 ttattatcac tttgggggtc ggcctgtgtc ccccgaaacg tctgcacctt ctgacgcagc 960
 44 ctgagaatga cctgcccctg ccccgagcct actctgtgta atagaataaa ggccctgcgtg 1020
 45 tgtctgtgga aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1059
 47 <210> SEQ ID NO: 2
 48 <211> LENGTH: 1461
 49 <212> TYPE: DNA
 50 <213> ORGANISM: Homo sapiens
 52 <220> FEATURE:
 53 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 55 <400> SEQUENCE: 2
 56 accccatctt catctggcct tgactctgcc cttgaggggc ctagggggtg agccagcctg 60
 57 ctccgagctc cctgcagat ggaggaggcc atcctggctc cctgcgtgct ggggctcctg 120
 58 ctgctgcoca tcctggccat gttgatggca ctgtgtgtgc actgccacag actgccaggc 180
 59 tcctacgaca gcacatctc agatagtgtt tatccaagg gcacccagtt caaacggcct 240
 60 cacacggttg cccctggccc acctgcctac ccacctgtca cctcctaccc acccctgagc 300

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/597,920
 DATE: 07/01/2000
 TIME: 12:40:32

Input Set : A:\05033.app
 Output Set: N:\CRF3\07012000\I597920.raw

```

61 cagccagacc tgcctcccat cccaagatcc ccgcagcccc ttgggggctc ccaccggacg 360
62 ccattcttcc cggcgggatt ctgatggtgc caacagtgtg gcgagctacg agaacgaggg 420
63 tgcgtctggg atccgagggtg cccaggtctg gtggggagtc tggggtccgt cctggactag 480
64 gctgaccctt gtgtcgttac ccccagaacc agcctgtgag gatgcagatg aggatgagga 540
65 cgactatcac aaccagggct acctgggtgt gcttctgac agcaccgccg ccactagcac 600
66 tgcctgcccc tcagctctct cactcagcac ccttggtatc cgagacagtg ccttctccat 660
67 ggagtccatt gatgattacg tgaacgttcc ggagagcggg gagagcgag aagcgtctct 720
68 ggaaggcagc cgggagtatg tgaatgtgtc ccaggaactg catcctggag cggctaagac 780
69 tgagccctgc gccctgagtt cccagggagg agaggaagtg gaggaagagg gggctccaga 840
70 ttacgagaat ctgcaggagc tgaactgagg gcctgtggag gccgagtcgt tccctggaacc 900
71 aggccttgctt gggacggctg agctgggcag ctggaagtgg ctctggggtc ctcacatggc 960
72 gtccctgcctt tgcctcagcc tgacaacagc ctgagaaatc ccccgtaac ttattatcac 1020
73 tttgggggtt gccctgtgtc ccccgaacgc tctgcacctt ctgacgcagc ctgagaatga 1080
74 cctgcctctg ccccgacctt actctgtgta atagaataaa ggctgcgtg tgtctgtgtt 1140
75 gagcgtgctg ctgtgtgtgc ctgtgtgcga gtctgagtc gagatttga gatgtctctg 1200
76 tgtgtttgtg tgtatctgtg ggtctccatc ctccatgggg gctcagccag gtgctgtgac 1260
77 acccccttct tgaatgaagc cttctgacct gggctggcac tgctgggggt gaggacacat 1320
78 tgccccatga gacagtccca gaacacggca gctgctggct gtgacaatgg tttcaccato 1380
79 cttagaccaa gggatgggac ctgatgacct gggaggactc ttttagttct tacctcttgt 1440
80 ggttctcaat aaaacagaaac g 1461
82 <210> SEQ ID NO: 3
83 <211> LENGTH: 1260
84 <212> TYPE: DNA
85 <213> ORGANISM: Mus musculus
87 <400> SEQUENCE: 3
88 ggcacgagca ggcggggagc aagaaagggg caggtacagc tgggcacggg gatcgtgcag 60
89 ctggtagctg ggcacgggc cccagctctg gctctggggc gagcaccttt ccagagccaa 120
90 cactgctctc aactcagtc agcaagagag gggagccatc cagccccgaa aggatacggc 180
91 tgctactgca cgggcgggac ccaggctgga gcccgcttgg tcccatatcc ctgctgccac 240
92 tctgtctcga ggggctgcag tgcagcaggg cctgtggcag gtgctctgca gatggaagca 300
93 gacgccttga gcccggtggg gctgggcctc ctgctgtctc ccttcttggg cagcctctctg 360
94 gctgcccctg gcgtgcgctg ccgtgagttg ccagtctcct atgacagcac ttccacagag 420
95 agtttgtacc caagaagcat cctcatcaag ccacctcaaa taaccgtccc ccgaacacct 480
96 gctgtttcct acctctagt cacttctctc ccacctctga ggcagccaga cctgctcccc 540
97 atcccagat cccacagacc ccttgggggt tcccatcgga tgccatcttc ccagcagaat 600
98 tcagatgatg ccaacagtgt ggcaagctac gagaaccagg agccagcctg taagaatgtg 660
99 gatgcagatg aggatgaaga cgactatccc aacggctacc tagtgggtgt gcctgacagt 720
100 agtccctgctg ccgtccctgt tgtctcctct gctcctgtgc ctgcaacccc tgaccttga 780
101 gacagtgcct tctctgtgga gtcgtgtgaa gattacgtga atgttctctg agtgaggag 840
102 agcgcagagg cgtctctgga tgggagccgg gagtatgtga atgtgtcccc agagcagcag 900
103 ccagtaccca gggctgagct ggctctctgt aactcccagg aggtggaaga cgaaggagaa 960
104 gaggaagggg tggatggaga ggaagctccc gactatgaga atctacagga gcttaactga 1020
105 aagcctactg cagctgtctg tcctgaaact ggacttgcgt ggggtgtcgt aagaggatcc 1080
106 catttgatct ctgccttgcc acagcctgag aatcttcccc taacttattg tcaacttggg 1140
107 gtccagtctg tgcctccaat attctgtacc ttctgataaa gcctgagaat gaatctgggt 1200
108 ccagccagac catgtcatgg aataaaggcc atgtgacata aaaaaaaaaa aaaaaaaaaa 1260
110 <210> SEQ ID NO: 4
111 <211> LENGTH: 233
112 <212> TYPE: PRT

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/597,920
 DATE: 07/01/2000
 TIME: 12:40:32

Input Set : A:\05033.app
 Output Set: N:\CRF3\07012000\I597920.raw

```

113 <213> ORGANISM: Homo sapiens
115 <400> SEQUENCE: 4
116 Met Glu Glu Ala Ile Leu Val Pro Cys Val Leu Gly Leu Leu Leu Leu
117 1 5 10 15
119 Pro Ile Leu Ala Met Leu Met Ala Leu Cys Val His Cys His Arg Leu
120 20 25 30
122 Pro Gly Ser Tyr Asp Ser Thr Ser Ser Asp Ser Leu Tyr Pro Arg Gly
123 35 40 45
125 Ile Gln Phe Lys Arg Pro His Thr Val Ala Pro Trp Pro Pro Ala Tyr
126 50 55 60
128 Pro Pro Val Thr Ser Tyr Pro Pro Leu Ser Gln Pro Asp Leu Leu Pro
129 65 70 75 80
131 Ile Pro Arg Ser Pro Gln Pro Leu Gly Gly Ser His Arg Thr Pro Ser
132 85 90 95
134 Ser Arg Arg Asp Ser Asp Gly Ala Asn Ser Val Ala Ser Tyr Glu Asn
135 100 105 110
137 Glu Glu Pro Ala Cys Glu Asp Ala Asp Glu Asp Glu Asp Tyr His
138 115 120 125
140 Asn Pro Gly Tyr Leu Val Val Leu Pro Asp Ser Thr Pro Ala Thr Ser
141 130 135 140
143 Thr Ala Ala Pro Ser Ala Pro Ala Leu Ser Thr Pro Gly Ile Arg Asp
144 145 150 155 160
146 Ser Ala Phe Ser Met Glu Ser Ile Asp Asp Tyr Val Asn Val Pro Glu
147 165 170 175
149 Ser Gly Glu Ser Ala Glu Ala Ser Leu Asp Gly Ser Arg Glu Tyr Val
150 180 185 190
152 Asn Val Ser Gln Glu Leu His Pro Gly Ala Ala Lys Thr Glu Pro Ala
153 195 200 205
155 Ala Leu Ser Ser Gln Glu Ala Glu Glu Val Glu Glu Gly Ala Pro
156 210 215 220
158 Asp Tyr Glu Asn Leu Gln Glu Leu Asn
159 225 230
162 <210> SEQ ID NO: 5
163 <211> LENGTH: 242
164 <212> TYPE: PRT
165 <213> ORGANISM: Mus musculus
167 <400> SEQUENCE: 5
168 Met Glu Ala Asp Ala Leu Ser Pro Val Gly Leu Gly Leu Leu Leu Leu
169 1 5 10 15
171 Pro Phe Leu Val Thr Leu Leu Ala Ala Leu Cys Val Arg Cys Arg Glu
172 20 25 30
174 Leu Pro Val Ser Tyr Asp Ser Thr Ser Thr Glu Ser Leu Tyr Pro Arg
175 35 40 45
177 Ser Ile Leu Ile Lys Pro Pro Gln Ile Thr Val Pro Arg Thr Pro Ala
178 50 55 60
180 Val Ser Tyr Pro Leu Val Thr Ser Phe Pro Pro Leu Arg Gln Pro Asp
181 65 70 75 80
183 Leu Leu Pro Ile Pro Arg Ser Pro Gln Pro Leu Gly Gly Ser His Arg
184 85 90 95

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/597,920
 DATE: 07/01/2000
 TIME: 12:40:32

Input Set : A:\05033.app
 Output Set: N:\CRF3\07012000\I597920.raw

```

186 Met Pro Ser Ser Gln Gln Asn Ser Asp Asp Ala Asn Ser Val Ala Ser
187      100      105      110
189 Tyr Glu Asn Gln Glu Pro Ala Cys Lys Asn Val Asp Ala Asp Glu Asp
190      115      120      125
192 Glu Asp Asp Tyr Pro Asn Gly Tyr Leu Val Val Leu Pro Asp Ser Ser
193      130      135      140
195 Pro Ala Ala Val Pro Val Val Ser Ser Ala Pro Val Pro Ser Asn Pro
196 145      150      155      160
198 Asp Leu Gly Asp Ser Ala Phe Ser Val Glu Ser Cys Glu Asp Tyr Val
199      165      170      175
201 Asn Val Pro Glu Ser Glu Glu Ser Ala Glu Ala Ser Leu Asp Gly Ser
202      180      185      190
204 Arg Glu Tyr Val Asn Val Ser Pro Glu Gln Gln Pro Val Thr Arg Ala
205      195      200      205
207 Glu Leu Ala Ser Val Asn Ser Ser Gln Glu Val Glu Asp Glu Gly Glu Glu
208      210      215      220
210 Glu Gly Val Asp Gly Glu Glu Ala Pro Asp Tyr Glu Asn Leu Gln Glu
211 225      230      235      240
213 Leu Asn
217 <210> SEQ ID NO: 6
218 <211> LENGTH: 2450
219 <212> TYPE: DNA
220 <213> ORGANISM: Homo sapiens
222 <400> SEQUENCE: 6
223 ggaatagggt agtttcagac aagcctgctt gccggagctc agcagacacc aggccttcctg 60
224 ggcaggcctg gccaccctg ggcctcagag ctgctgctgg gccattcaga accggctctc 120
225 cattggcatt gggaccagag accccgcaag tggcctgttt gcctggacat ccacctgtac 180
226 gtccccaggt ttccggaggc ccagggggcg tgccagacc cgcggcgcac ctgcccttct 240
227 tctacggcag catctcgctt gccgaggccg aggagcacct gaagctggcg ggcattggcg 300
228 acgggctctt cctgctgcgc cagtgcctgc gctcgtggg cggtatgtg ctgtcgtctg 360
229 tgcacgatgt gcgttccac cactttccca tcgagcgcca gctcaacggc acctacgcca 420
230 ttgccggcgg caaagcgcac tgtggaccgg cagagctctg cgagttctac tcgcgcgacc 480
231 ccgacgggct gccctgcaac ctgcgcaagc cgtgcaaccg gccgtcgggc ctcgagccgc 540
232 agccgggggt ctcgactgc ctgcgagacg ccatggtgctg tgactacgtg cgccagacgt 600
233 ggaagctgga gggcgaggcc ctggagcagg ccatcatcag ccaggccccg cagggtggaga 660
234 agctcattgc tacgacggcc cagagcgga tgccctggta ccacagcagc ctgacgcgtg 720
235 aggagggcga gcgcaaactt tactctgggg cgcagaccga cggcaagtgc ctgctgaggc 780
236 cgcggaagga gcgggcacat acgccctgtc cctcatctat ggggaagacg tgtaccacta 840
237 cctcatcagc caagacaagg cgggcaagta ctgcattccc gagggcacca agtttgacac 900
238 gctctggcag ctggtggagt atctgaagct gaaggcggac gggctcatct actgectgaa 960
239 ggaggcctgc cccaacagca gtgccagcaa cgcctcaggg gctgctgctc ccacactccc 1020
240 agcccaccca tccacgttga ctcatctca gagacgaatc gacaccctca actcagatgg 1080
241 atacaccctg agccagcacg cataacgtcc ccagacaaac cgcggccgat gccatggac 1140
242 acgagcgtgt atgagagccc ctacagcgac ccagaggagc tcaaggacaa gaagctcttc 1200
243 ctgaagcgcg ataacctcct catagctgac attgaacttg gctgcgccaa ctttggtca 1260
244 gtgcgcccag gcgtgtaccg catgcgcaag aagcagatcg acgtggccat caaggtgctg 1320
245 aagcagggca cggagaaggc agacacggaa gagatgatgc gcgagggcga gatcatgcac 1380
246 cagctggaca accctacat cgtgcggctc attggcgtct gccaggccga ggccctcatg 1440
247 ctggtcatgg agatggctgg ggcggggccg ctgcacaagt tcctggtcgg caagagggag 1500

```

RAW SEQUENCE LISTING

DATE: 07/01/2000

PATENT APPLICATION: US/09/597,920

TIME: 12:40:32

Input Set : A:\05033.app

Output Set: N:\CRF3\07012000\I597920.raw

```

248 gagatccctg tgagcaatgt ggccgagctg ctgcaccagg tgtccatggg gatgaagtac 1560
249 ctggaggaga agaactttgt gcaccgtgac ctggcggccc gcaacgtcct gctgggttaa 1620
250 ccggcactac gccaatgaca gcgacttggc ctctccaaag cactgggtgc cgacgacagc 1680
251 tactacactg ccgctcagc agggaaagtgg ccgctcaagt ggtacgcacc cgaatgcac 1740
252 aacttcgcga agttctccag ccgcagcgat gtctggagct atggggtcac catgtgggag 1800
253 gccttgctct acggccagaa gccctacaag aagatgaaag ggccggaggt catggccttc 1860
254 atcgagcagg gcaagcggat ggagtgccca ccagagtgtc caccggaact gtacgcactc 1920
255 atgagtgact gctggatcta caagtgggag gatcgccccg acttctgac cgtggagcag 1980
256 cgcatacgag cctgttacta cagcctggcc agcaagggtg aagggccccc aggcagcaca 2040
257 cagaaggctg aggtgcctg tgcctgagct ccgctgccc aggggagccc tcacgcgg 2100
258 ctcttcccca cctcagccc caccagagt cctgcagtct ggtgagccc tgcctgggtg 2160
259 tctccacaca cagctgggct gtgtagggg gtgtctcagg ccacaccggc cttgcattgc 2220
260 ctgctggccc cctgtctct tctggtggg gagcagggag gtccgggagg gtgcgctgt 2280
261 gcagcctgtc ctgggctggt ggctcccgga gggccctgag ctgagggcat tgcctacag 2340
262 gatgccttcc cctgggccct gacattggag cctgggcata ctcaggtggt caggcgtaga 2400
263 tcaccagaat aaaccagct tccctcttga aaaaaaaaaa aaaaaaac 2450
265 <210> SEQ ID NO: 7
266 <211> LENGTH: 617
267 <212> TYPE: PRT
268 <213> ORGANISM: Homo sapiens
270 <400> SEQUENCE: 7
271 Met Pro Asp Pro Ala Ala His Leu Pro Phe Phe Tyr Gly Ser Ile Ser
272 1 5 10 15
274 Arg Ala Glu Ala Glu His Leu Lys Leu Ala Gly Met Ala Asp Gly
275 20 25 30
277 Leu Phe Leu Leu Arg Gln Cys Leu Arg Ser Leu Gly Gly Tyr Val Leu
278 35 40 45
280 Ser Leu Val Asp Asp Val Arg Phe His His Phe Pro Ile Glu Arg Gln
281 50 55 60
283 Leu Asn Gly Thr Tyr Ala Ile Ala Gly Gly Lys Ala His Cys Gly Pro
284 65 70 75 80
286 Ala Glu Leu Cys Gln Phe Tyr Ser Gln Asp Pro Asp Gly Leu Pro Cys
287 85 90 95
289 Asn Leu Arg Asn Ala Cys Asn Arg Pro Pro Gly Leu Glu Pro Gln Pro
290 100 105 110
292 Gly Val Phe Asp Cys Leu Arg Asp Ala Met Val Arg Asp Tyr Val Arg
293 115 120 125
295 Gln Thr Trp Lys Leu Glu Gly Asp Ala Leu Glu Gln Ala Ile Ile Ser
296 130 135 140
298 Gln Ala Pro Gln Val Glu Lys Leu Ile Ala Thr Thr Ala His Glu Arg
299 145 150 155 160
301 Met Pro Trp Tyr His Ser Ser Leu Thr Arg Glu Glu Ala Glu Arg Lys
302 165 170 175
304 Leu Tyr Ser Gly Gln Gln Thr Asp Gly Lys Phe Leu Leu Arg Pro Arg
305 180 185 190
307 Lys Glu Gln Gly Thr Tyr Ala Leu Ser Leu Val Tyr Gly Lys Thr Val
308 195 200 205
310 Tyr His Tyr Leu Ile Ser Gln Asp Lys Ala Gly Lys Tyr Cys Ile Pro
311 210 215 220

```

VERIFICATION SUMMARY

DATE: 07/01/2000

PATENT APPLICATION: US/09/597,920

TIME: 12:40:33

Input Set : A:\05033.app

Output Set: N:\CRF3\07012000\I597920.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date